

(19) World Intellectual Property  
Organization  
International Bureau



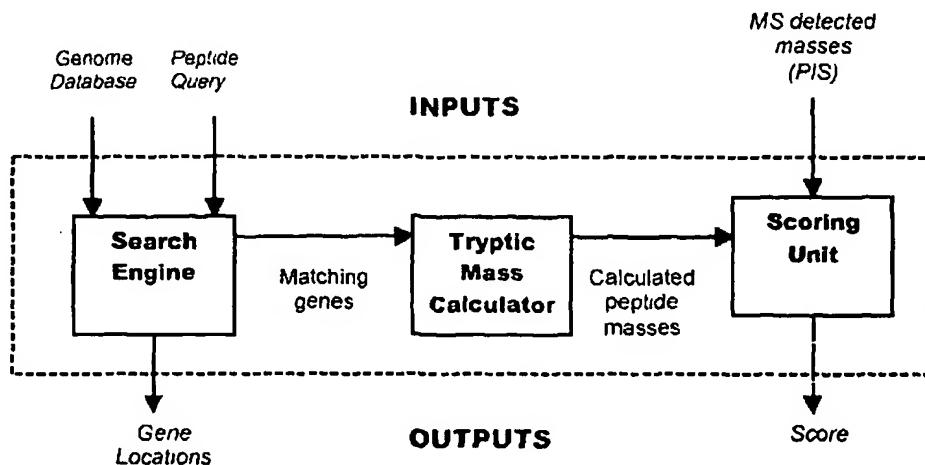
(43) International Publication Date  
16 December 2004 (16.12.2004)

PCT

(10) International Publication Number  
**WO 2004/109579 A3**

- (51) International Patent Classification<sup>7</sup>: **G06F 19/00**
- (21) International Application Number:  
PCT/CA2004/000853
- (22) International Filing Date: 9 June 2004 (09.06.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:  
60/477,076 9 June 2003 (09.06.2003) US
- (71) Applicant (for all designated States except US): **MOUNT SINAI HOSPITAL [CA/CA]**; 600 University Avenue, Toronto, Ontario M5G1X5 (CA).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **HOGUE, Christopher [CA/CA]**; 73 Richmond Street, Suite 212, Toronto, Ontario M5H 4E8 (CA). **ROSE, Jonathan [CA/CA]**; 181 Albany Avenue, Toronto, Ontario M5R 3C7 (CA). **ALEX, Anish [CA/CA]**; 90 Trailridge Crescent, Toronto, Ontario M1E 4C5 (CA).
- (74) Agents: **KURDYDYK, Linda et al.; McCarthy Tetrault**, 66 Wellington Street West, Suite 4700, P.O. Box 48, Toronto Dominion Centre, Toronto, Ontario M5K 1E6 (CA).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:  
— with international search report
- (88) Date of publication of the international search report:  
2 June 2005
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROTEIN IDENTIFICATION METHODS AND SYSTEMS



(57) Abstract: The present invention relates to methods and systems for identifying proteins. In particular the invention provides a method for identifying a protein through amino acid sequences of one or more query peptides generated from the protein. The method involves translating amino acid sequences of the query peptides to all possible codons from which the peptides can be synthesized to prepare strings of codons. Known nucleic acid sequences, in particular a set of known nucleic acid sequences including a genome, are searched to locate one or more known nucleic acids that comprise regions that match the strings of codons. Matching nucleic acids are ranked to identify nucleic acids that are true coding regions for the protein to thereby identify the protein.